

On a representation of partially-distinguishable populations

Jeremie Houssineau^{a,*}, Daniel E. Clark^b

^a*Department of Statistics and Applied Probability, National University of Singapore, Singapore, 117546.*

^b*School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, UK.*

Abstract

A way of representing heterogeneous stochastic populations that are composed of sub-populations with different levels of distinguishability is introduced together with an analysis of its properties. In particular, it is demonstrated that any instance of this representation where individuals are independent can be related to a point process on the set of probability measures on the individual state space. The introduction of the proposed representation is fully constructive which guarantees the meaningfulness of the approach.

Keywords: Stochastic population, Point process

2010 MSC: 60A10, 62C10

Introduction

Stochastic populations such as probabilistic multi-object systems are of central importance in many areas within systems biology [1], robotics [2] or computer vision [3]. In some cases, the sole interest is in their global characteristics, such as when only their cardinality is studied, e.g. in population dynamics [4, 5], or when spatial information is meant to be unspecific, as with point processes [6, 7]. In some other cases, all the individuals of the population can be clearly identified and the way the population is represented becomes less fundamental since the problem can be recast into a collection of individual-wise representations. Except in these specific cases, the representation of stochastic populations remains mostly unexplored, in spite of their ubiquity. In general, the population might be only partially distinguishable, i.e. some individuals might be identified while another sub-population might only be described by unspecific representations, e.g. by its cardinality. The objective in this article is to find a natural way of representing these stochastic partially-distinguishable populations. The

*Corresponding author

Email addresses: `houssineau.j@gmail.com` (Jeremie Houssineau), `d.e.clark@hw.ac.uk` (Daniel E. Clark)

underlying motivation is that a natural representation should not only be useful in theory when expressing different results and properties, but also in practice when devising approximation algorithms for the induced probability laws. Figure 1 shows examples of samples drawn for distributions with different degrees of distinguishability, hinting at the possible drawbacks of using indistinguishable representations for distinguishable populations.

Example 1. In the context of Bayesian data assimilation for stochastic populations, sub-populations that have never been observed are often well modelled by indistinguishable representations, e.g. if individuals live on the real line and if new individuals are known to appear either at point $a \in \mathbb{R}$ or at point $b \in \mathbb{R}$ different from a , then it is not unnatural that individuals might appear at the same point, either a or b . However, if one individual has been observed at a and another one at b , then using a representation that allows these two individuals to be both at either a or b would often be inappropriate. Overall, different sub-population require different levels of distinguishability and a suitable stochastic representation should be able to deal with this modelling aspect.

One of the main application areas for the type of representation introduced in this article is in the Engineering discipline called multi-target tracking [8, 9], see e.g. [10, 11, 12] or [13, Chapt. 6] for a point-process-based formulation and analysis. In this context, the limitation of point processes is found in their inability to represent and propagate specific information about targets, or *tracks*. Since this is often the objective, heuristics are usually applied to the output of the point-process-based algorithm in order to produce tracks. However, since tracks themselves are often not only displayed to the operator but also used for further processing steps, the addition of an ad-hoc step at this stage of the algorithm prevents from performing these steps in a principled and integrated way. For instance, specific data assimilation [14] based on the proposed representation can be easily extended to include classification [15] or sensor management [16]. Existing applications of the proposed approach include space situational awareness [17], harbour surveillance [15] as well as multi-target tracking from radar data [18].

In order to build a natural representation of stochastic populations, it is convenient to start with an idealistic case in which the notion of partial distinguishability can be formalised, and so is done in Section 1. The concepts and notations introduced in Section 1 are then used as a basis for the introduction of a full representation in Section 2. An alternative formulation is finally introduced in Section 3, where simplifications are made in order to make the representation more practical.

Throughout the article, random variables will be implicitly assumed to be defined on the complete probability space $(\Omega, \Sigma, \mathbb{P})$. For any set A , denote $\Pi(A)$ the set of equivalence relations on A , and denote O and I the minimal and maximal equivalence relations respectively, i.e. xOy is false and xIy is true for any $x, y \in A$.

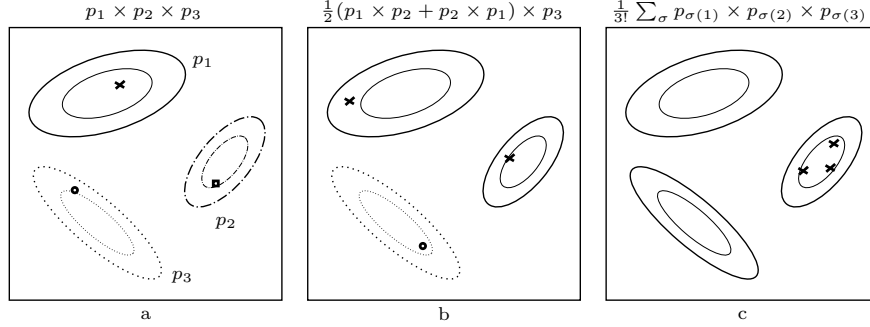


Figure 1: Distributions and samples for different degrees of distinguishability: (a) fully distinguishable, (b) partially distinguishable and (c) indistinguishable. The samples in (c) show a drawback of using indistinguishable representations for distinguishable populations, i.e. there is no guarantee that the individual samples will come from different modes, as opposed to (a)

1. Describing a population

We consider a *representative* set \mathcal{X}_a , i.e. a set in which individuals of interest can be uniquely characterised. Because of this characterisation, a population, which can be intuitively understood as a collection of individuals, is formally defined as a subset of \mathcal{X}_a . The set \mathcal{X} of all possible populations is then defined as the set of all countable subsets of \mathcal{X}_a . In this way, the set \mathcal{X} is itself a representative set for populations.

An important aspect is that in practice, a more realistic set \mathbf{X} needs to be considered for the representation of individuals. This set is seen as being a projection of the set \mathcal{X}_a and we define $\phi : \mathcal{X}_a \rightarrow \mathbf{X}$ as the associated projection map. Such a simplification is required for most of the applications since the full characterisation of an individual is not usually considered accessible. For instance, the observation might not account for the shape, mass or composition of a given solid, so that only its centre of mass/volume can be inferred. One of the consequences of this simplified representation is that individuals might have the same state in \mathbf{X} . In the context of point process theory [19], processes that never have two individuals at the same point are called *simple*. Borrowing this term, we can impose that representations should not require *simplicity* in \mathbf{X} in general. A practical example of the meaning of the sets introduced so far is given in Figure 2.

The aptitude to obtain specific information, or *observability*, might not be sufficient to tell some of the individuals apart. Individuals that are in this situation are said to be *strongly indistinguishable*, i.e. they cannot be distinguished in their current states even with the best possible sources of information. Strongly indistinguishable individuals can be related through a relation $\tau \in \mathbf{\Pi}(\mathcal{X})$ defined as follows: two individuals $x, x' \in \mathcal{X}$ are strongly indistinguishable if and only if $x\tau x'$ holds. The set

$$\mathcal{Y} \doteq \{(\mathcal{X}, \tau) \text{ s.t. } \mathcal{X} \in \mathcal{X}, \tau \in \mathbf{\Pi}(\mathcal{X})\}$$

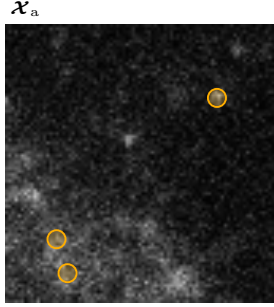


Figure 2: Image of proteins obtained by Fluorescence Microscopy, assumed to be the representative set \mathcal{X}_a in this context, where the 3 individual proteins are highlighted in by small circles constitute an example of population $\mathcal{X} \in \mathcal{X}$. The set \mathbf{X} might for instance describe only the position of the proteins in the image.

is introduced in order to represent partially-indistinguishable populations. When individuals are not strongly indistinguishable, they are said to be *weakly distinguishable*. Even when some individuals are weakly distinguishable, it could happen that the available information is not sufficient to tell them apart. We then say that these individuals are *weakly indistinguishable*. This concept clearly depends on the knowledge about the population and might evolve if additional information is made available. To sum up, strong indistinguishability is a state-dependent concept while weak indistinguishability is a probabilistic concept.

The description of the uncertainty on a given population $\mathcal{X} \in \mathcal{X}$ can be performed by associating every individual in \mathcal{X} with a random variable on \mathbf{X} . This solution, however, does not describe the relation between the different distributions related to different individuals, in particular with strongly indistinguishable ones. A global representation of uncertainty is thus sought. One of the most usual ways of describing multiple spatial entities as a whole is given by the theory of point processes. However, this theory is built on the following principle:

“We talk of the probability of finding a given number k of points in a set A : we do not give names to the individual points and ask for the probability of finding k specified individuals within the set A . Nevertheless, this latter approach is quite possible (indeed, natural) in contexts where the points refer to individual particles, animals, plants and so on.” [19, p. 124]

Yet, we wish to model the partially-indistinguishable nature of the individuals in \mathcal{X} without assuming that they are all strongly indistinguishable, i.e. without assuming that $\tau = I$. The study of populations composed of indistinguishable individuals is already challenging due to the difficulty in finding a consistent way of describing multiple individuals within a single stochastic object. Examples of questions arising from this issue are: Should the individuals be ordered even though there is no natural way of defining the order? Should

the individuals be assumed to be represented at different points of the state space in order to enable a set representation? Should the population be assumed finite in order to proceed to the analysis? There are different ways of answering these questions and each way has to be proved equivalent in some sense to the others [20, 21]. The representation of partially indistinguishable populations raises many additional and equivalently difficult questions. Alternative representations of stochastic populations have to be found in order to tackle this issue.

2. Representing a population

Based on the set \mathcal{X} of all possible populations and on the set \mathbf{X} on which all individuals are represented, we describe a versatile way of introducing randomness in the states of the individuals in \mathbf{X} which conveys the concept of strong indistinguishability. This is first achieved for a fixed population in Section 2.1 before tackling the full generality of the problem in Section 2.2.

2.1. For a given population

We assume that the set \mathbf{X} can be written as the union of an Euclidean space \mathbf{X}^\bullet and an *isolated point* ψ . The latter can be viewed as an *empty state* and is used to provide an image to individuals that cannot be represented on \mathbf{X}^\bullet such as individuals that are outside of the zone of interest.

2.1.1. Construction

Let $\mathcal{Y} = (\mathcal{X}, \tau) \in \mathcal{Y}$ be a partially-distinguishable population of interest, i.e. a set \mathcal{X} of individuals characterised in \mathcal{X}_a that is equipped with an equivalence relation τ connecting strongly indistinguishable individuals. The objective is to include the relation between the individuals of \mathcal{X} in the probabilistic modelling of the population. We first introduce the set

$$\mathbf{F}_{\mathcal{Y}} \doteq \{f : \mathcal{X} \rightarrow \mathbf{X} \text{ s.t. } |f^{-1}[\mathbf{X}^\bullet]| < \infty\}$$

that is composed of mappings $f : \mathcal{X} \rightarrow \mathbf{X}$ that map finitely many individuals to \mathbf{X}^\bullet . This condition facilitates the definition of various types of operations on individuals but can be relaxed without inducing major changes in the following results. The set \mathcal{X} is used as a way of indexing the states in \mathbf{X} and the actual knowledge of the full individual characteristics $x \in \mathcal{X}$ is not used. Otherwise, the state of an individual $x \in \mathcal{X}$ could be directly obtained from the projection $\phi(x) \in \mathbf{X}$. At the end of this section, we will derive a formulation that ensures that \mathcal{X} cannot be used to hold information on the state of individuals.

A suitable σ -algebra of subsets of $\mathbf{F}_{\mathcal{Y}}$, denoted $\mathcal{F}_{\mathcal{Y}}^*$ can be introduced as follows: There is a natural topology on $\mathbf{F}_{\mathcal{Y}}$ that is generated by open sets of the same form as

$$A = \{f \text{ s.t. } (\forall x \in \mathcal{X}) \ f(x) \in A_x\},$$

where A_x is an open set in \mathbf{X} that differs from $\{\psi\}$ for finitely many $x \in \mathcal{X}$ only. Note that $\{\psi\}$ is indeed open as an isolated point. This topology is denoted $\mathcal{T}_{\mathcal{Y}}^*$

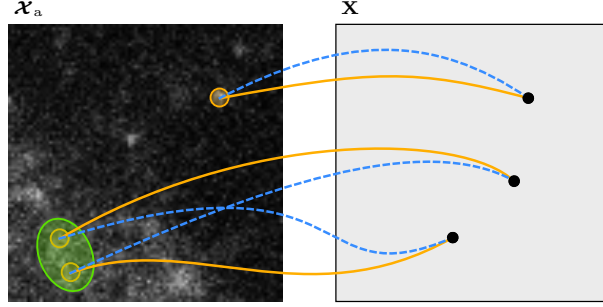


Figure 3: The two individuals in the bottom left corner are assumed to be indistinguishable so that the two displayed maps should be considered as equivalent since it should not be possible to obtain the states of indistinguishable individuals specifically.

and $\mathcal{F}_{\mathcal{Y}}^*$ is defined as the corresponding Borel σ -algebra. Representations of the population \mathcal{X} can thus be given as random variables in the measurable space of mappings $(\mathbf{F}_{\mathcal{Y}}, \mathcal{F}_{\mathcal{Y}}^*)$.

A random variable \mathfrak{F} on $(\mathbf{F}_{\mathcal{Y}}, \mathcal{F}_{\mathcal{Y}}^*)$ represents all the individuals in \mathcal{X} on \mathbf{X} and is equivalent to a collection of possibly correlated random variables, since indistinguishability has not been taken into account yet.

When two individuals in \mathcal{X} are strongly indistinguishable, we expect that individual characterisations would not be available, even when considering a specific outcome $\omega \in \Omega$. Random variables on $(\mathbf{F}_{\mathcal{Y}}, \mathcal{F}_{\mathcal{Y}}^*)$ that do not respect this constraint would be mistakenly distinguishing individuals that are strongly indistinguishable, as shown in Figure 3. The space $(\mathbf{F}_{\mathcal{Y}}, \mathcal{F}_{\mathcal{Y}}^*)$ is then not fully satisfying as it does not ensure that indistinguishable individuals are well represented.

A natural way of circumventing this incomplete representation of the structured population \mathcal{Y} is to make the σ -algebra $\mathcal{F}_{\mathcal{Y}}^*$ coarser by “gluing” together functions that distinguish indistinguishable individuals.

Example 2. Suppose that $\mathcal{Y} = (\{x, x'\}, I)$, i.e. \mathcal{Y} is made of two indistinguishable individuals so that $\mathcal{X}/\tau = \{\{x, x'\}\}$. Additionally suppose that $\mathbf{X} = \mathbf{X}^\bullet = \{\mathbf{x}, \mathbf{x}'\}$, i.e. there are only two possible states for the individuals x and x' , and assume that \mathbf{X} is also representative so that x and x' must have different states in \mathbf{X} . There are only $2! = 2$ different distinguishable outcomes f, g in $\mathbf{F}_{\mathcal{Y}}$ defined by their respective graph as $\{(x, \mathbf{x}), (x', \mathbf{x}')\}$ and $\{(x, \mathbf{x}'), (x', \mathbf{x})\}$. To ensure that the individuals x, x' are indistinguishable, one can glue together these two symmetrical outcomes and define a new set of functions as $\{\{f, g\}\}$ (note the additional curly brackets). There is now only one outcome $\{f, g\}$ that does not allow for distinguishing the individuals x and x' as required.

Following Example 2 and denoting $\text{Sym}(\mathcal{X}, \tau)$ the subgroup of permutations on \mathcal{X} agreeing with the equivalence relation τ , i.e. the ones permuting indistinguishable individuals only, we introduce a binary relation on $\mathbf{F}_{\mathcal{Y}}$ as follows.

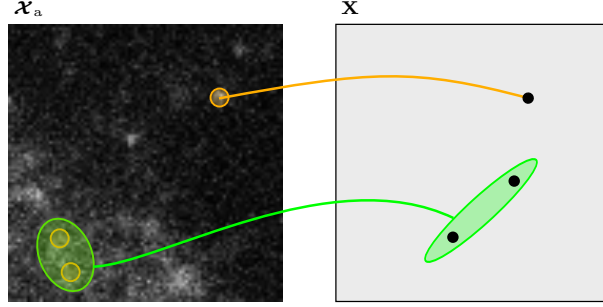


Figure 4: Representation of the element of the quotient space \mathbf{F}_Y/ρ associated to the case displayed in Figure 3. The specific states of indistinguishable individuals are no longer accessible as required.

Definition 1. A binary relation ρ on \mathbf{F}_Y is said to be induced by the equivalence relation τ if it holds that

$$(\forall f, f' \in \mathbf{F}_Y) \quad f \rho f' \Leftrightarrow \exists \sigma \in \text{Sym}(\mathcal{X}, \tau) (f = f' \circ \sigma). \quad (1)$$

Intuitively, elements of \mathbf{F}_Y are related through a binary relation whenever they only differ by a permutation of indistinguishable individuals. A representation of the elements of the quotient space \mathbf{F}_Y/ρ is given in Figure 4. This binary relation can be proved to have additional properties.

Proposition 1. *The equivalence relation τ induces a unique binary relation on $\mathbf{F}_Y(\mathbf{X})$, and this binary relation is an equivalence relation.*

The proof of Proposition 1 relies mostly on the group nature of $\text{Sym}(\mathcal{X}, \tau)$, as a subgroup of $\text{Sym}(\mathcal{X})$. Consequently, only the specific group properties of $\text{Sym}(\mathcal{X}, \tau)$ will be invoked when proving that the induced binary relation is an equivalence relation.

Proof. (Uniqueness) Let ρ and ρ' be two binary relations induced by τ . We want to prove that $f \rho f' \Leftrightarrow f \rho' f'$ holds for any $f, f' \in \mathbf{F}_Y(\mathbf{X})$. Let σ, σ' be the two permutations in $\text{Sym}(\mathcal{X}, \tau)$ satisfying (1) for ρ and ρ' respectively. There exists σ'' in $\text{Sym}(\mathcal{X}, \tau)$ such that $\sigma \circ \sigma'' = \sigma'$, proving the uniqueness.

(Reflexivity) The identity is in $\text{Sym}(\mathcal{X}, \tau)$.

(Symmetry) Existence of an inverse element in $\text{Sym}(\mathcal{X}, \tau)$.

(Transitivity) Closure of $\text{Sym}(\mathcal{X}, \tau)$. □

Let ρ denote the unique equivalence relation on \mathbf{F}_Y induced by τ and let ξ_ρ be the quotient map from \mathbf{F}_Y to \mathbf{F}_Y/ρ induced by ρ . We introduce a σ -algebra of subsets of \mathbf{F}_Y , denoted \mathcal{F}_Y , which does not allow for distinguishing strongly indistinguishable individuals: Let \mathcal{T}_Y denote the initial topology on \mathbf{F}_Y induced by the quotient map ξ_ρ . We can verify that $\mathcal{T}_Y \subseteq \mathcal{T}_Y^*$ holds, meaning that there are fewer open subsets in \mathcal{T}_Y when compared to \mathcal{T}_Y^* . The Borel σ -algebra induced

by $\mathcal{T}_{\mathcal{Y}}$ is denoted $\mathcal{F}_{\mathcal{Y}}$. A reference measure on $(\mathbf{F}_{\mathcal{Y}}, \mathcal{F}_{\mathcal{Y}})$ can be easily deduced from the reference measure on \mathbf{X} , e.g. the Lebesgue measure. Random variables on $(\mathbf{F}_{\mathcal{Y}}, \mathcal{F}_{\mathcal{Y}})$ characterise subsets of indistinguishable individuals rather than individuals themselves, as required.

2.1.2. Independence and weak indistinguishability

Now equipped with suitable spaces for considering the representation of partially-indistinguishable populations, we study the properties of probability measures on $(\mathbf{F}_{\mathcal{Y}}, \mathcal{F}_{\mathcal{Y}})$. Since populations have an intrinsic multivariate nature, it is natural to introduce a notion of independence for probability measures on $\mathbf{F}_{\mathcal{Y}}$ as in the following definition.

Definition 2. The individuals in \mathcal{X} are said to be *independent* if the law P on $\mathbf{F}_{\mathcal{Y}}$ verifies

$$P(F) = \int \mathbf{1}_F(x \mapsto \mathbf{y}_x) \prod_{x \in \mathcal{X}} p_x(d\mathbf{y}_x) \quad (2)$$

for any $F \in \mathcal{F}_{\mathcal{Y}}$, where $\{p_x\}_{x \in \mathcal{X}}$ a family of probability measures on \mathbf{X} .

The expression (2) of Definition 2 is a convolution of measures based on the operation of creating a function in $\mathbf{X}^{\mathcal{X}}$ out of a value in \mathbf{X} for each the individuals in \mathcal{X} . This notion of independence will be useful as an example of concepts and operations that will be defined in the general case.

The notion of weak indistinguishability that was introduced in Section 1 has not been translated into practical terms yet. As opposed to strongly indistinguishable individuals that are bound through the events in $\mathcal{F}_{\mathcal{Y}}$, it just happens that there is no specific knowledge about weakly indistinguishable individuals. As a result, weak indistinguishability is a fully probabilistic concept. In order to formally define it, we introduce a mapping T_{σ} from $\mathbf{F}_{\mathcal{Y}}$ into itself for any given $\sigma \in \text{Sym}(\mathcal{X})$ defined by

$$T_{\sigma} : f \mapsto f \circ \sigma. \quad (3)$$

Mappings of this form describe the changes induced by swapping individuals. It is therefore suitable for expressing properties of symmetry for probability measures as in the following definition.

Definition 3. Let P be a probability measure on $\mathbf{F}_{\mathcal{Y}}$. The relation of weak indistinguishability induced by P on \mathcal{X} is defined as

$$\eta = \sup \{ \eta' \in \mathbf{\Pi}(\mathcal{X}) \text{ s.t. } (\forall \sigma \in \text{Sym}(\mathcal{X}, \eta')) \ P = (T_{\sigma})_* P \},$$

where $(T_{\sigma})_* P$ is the pushforward of P by the measurable mapping T_{σ} .

The relation of weak indistinguishability is an equivalence relation by definition. Since $\mathbf{\Pi}(\mathcal{X})$ is only a partially ordered set, the greatest element of a given subset might not exist, but it is necessarily unique if it exists. We can show that the relation η of weak indistinguishability exists by verifying that any

element $\eta' \neq \eta$ in the considered subset can only identify less symmetries than η . In other words, denoting $\Pi(\eta)$ the partition of \mathcal{X} induced by η , there exist at least two subsets in $\Pi(\eta')$ which union is a subset of $\Pi(\eta)$ so that $\Pi(\eta') \leq \Pi(\eta)$ holds for any η' in the subset of $\mathbf{\Pi}(\mathcal{X})$ of interest. Some of the properties of the relation of weak indistinguishability are given here using the notations of Definition 3.

Proposition 2. *It holds that $\eta \geq \tau$.*

Proof. Sets in the σ -algebra $\mathcal{F}_{\mathcal{Y}}$ of subsets of $\mathbf{F}_{\mathcal{Y}}$ do not allow for distinguishing individuals related by τ . Thus, for any given $\sigma \in \text{Sym}(\mathcal{X}, \tau)$ and $F \in \mathcal{F}_{\mathcal{Y}}$, it holds that $f \circ \sigma \in F$ for any $f \in F$ so that $P = (T_{\sigma})_* P$ is always true by construction. As a result, the equivalence relation τ is always in the set of which η is the greatest element. \square

Example 3. Reusing the notations of Definition 2 and assuming that the individuals in \mathcal{X} are independent under P and that η is the relation of weak indistinguishability induced by P , then for any pair (x, x') of individuals in \mathcal{X} , it holds that

$$(x\eta x') \Leftrightarrow (p_x = p_{x'}).$$

The representation of strongly indistinguishable individuals by random variables on $(\mathbf{F}_{\mathcal{Y}}, \mathcal{F}_{\mathcal{Y}})$ can be considered as satisfactory. Yet, the true population \mathcal{Y} was supposed to be known so far, even though it is only used as an indexing set, this cannot be assumed in general. It is thus necessary to find a way of dealing with unknown populations.

2.2. Stochastic representation

It is natural to reuse the same mechanisms as before to bypass the necessity of knowing the true population when describing it, i.e. by defining an appropriate equivalence relation and working on the σ -algebras induced by the corresponding quotient spaces. However, we will see that the approach that seems the most natural at first does not lead to a satisfactory result. Nonetheless, this approach is detailed here as it motivates the introduction of a more advanced construction.

2.2.1. Naive attempt

The most natural way to extend the results of the previous section to unknown populations is to consider the union of the sets $\mathbf{F}_{\mathcal{Y}}$ and to simplify it using an equivalence relation as previously. Let the set \mathbf{F} be defined as

$$\mathbf{F} \doteq \bigcup_{\mathcal{Y} \in \mathcal{Y}} \mathbf{F}_{\mathcal{Y}}.$$

Definition 4. Let $\mathcal{Y}, \mathcal{Y}' \in \mathcal{Y}$ be two populations equipped with a relation of strong indistinguishability defined via $\mathcal{Y} \doteq (\mathcal{X}, \tau)$ and $\mathcal{Y}' \doteq (\mathcal{X}', \tau')$. The binary relations \sim on \mathcal{X} and \approx on \mathcal{Y} are defined as follows

$$\mathcal{X} \sim \mathcal{X}' \Leftrightarrow |\mathcal{X}| = |\mathcal{X}'| \quad \text{and} \quad \mathcal{Y} \approx \mathcal{Y}' \Leftrightarrow \exists \nu : \mathcal{Y} \xrightarrow{\sim} \mathcal{Y}',$$

where $\xleftrightarrow{\sim}$ indicates a relation-preserving bijection. Also, for any $f \in \mathbf{F}_{\mathcal{Y}}$ and any $f' \in \mathbf{F}_{\mathcal{Y}'}$, let the binary relation ρ^* on \mathbf{F} be defined as

$$f \rho^* f' \Leftrightarrow \exists \nu : \mathcal{Y} \xleftrightarrow{\sim} \mathcal{Y}' (f = f' \circ \nu). \quad (4)$$

It is easy to prove that the binary relations \sim , \approx and ρ^* on the respective sets \mathcal{X} , \mathcal{Y} and \mathbf{F} are equivalence relations. Note that the relation \sim on \mathcal{X} can be equivalently defined as

$$\mathcal{X} \sim \mathcal{X}' \Leftrightarrow \exists \nu : \mathcal{X} \leftrightarrow \mathcal{X}',$$

where \leftrightarrow indicates a bijection. This alternative definition highlights the parallel with the equivalence relations \approx and ρ^* also introduced in Definition 4.

Equivalence classes in \mathbf{F}/ρ^* do not allow for distinguishing functions that give the same values in \mathbf{X} and have different domains. As before, an appropriate σ -algebra \mathcal{F} of subsets of \mathbf{F} can be deduced from the quotient space \mathbf{F}/ρ^* . A first clue that the equivalence relation ρ^* is over-simplifying the space \mathbf{F} is that

$$\text{Sym}(\mathcal{X}, \tau) \subseteq \{\nu \text{ s.t. } \nu : (\mathcal{X}, \tau) \xleftrightarrow{\sim} (\mathcal{X}, \tau)\},$$

for any $(\mathcal{X}, \tau) \in \mathcal{Y}$, with the inclusion being strict for $\mathcal{X} \neq \emptyset$ and $\tau \neq I$; if $\tau = O$ for instance then \mathbf{F}/ρ^* will make all individuals indistinguishable although they were initially weakly distinguishable. We can still verify that the space $(\mathbf{F}, \mathcal{F})$ is suitable in cases where all the individuals are strongly indistinguishable by showing the relation between the subset

$$\mathbf{F}_I \doteq \bigcup_{\mathcal{X} \in \mathcal{X}} \mathbf{F}_{(\mathcal{X}, I)}$$

of \mathbf{F} endowed with the σ -algebra \mathcal{F}_I induced by \mathcal{F} and the set $\mathbf{N}(\mathbf{X})$ of integer-valued measures, or counting measures, on \mathbf{X} equipped with its Borel σ -algebra $\mathcal{N}(\mathbf{X})$. Such a relation will ensure that random variables on $(\mathbf{F}_I, \mathcal{F}_I)$ will be equivalent to point processes on \mathbf{X} as expected. In the next theorem, $\text{dom}(f)$ will denote the domain of a given function f .

Theorem 1. *The mapping ξ defined as*

$$\begin{aligned} \xi : \mathbf{F}_I &\rightarrow \mathbf{N}(\mathbf{X}) \\ f &\mapsto \sum_{x \in \text{dom}(f)} \delta_{f(x)}, \end{aligned}$$

is $\mathcal{F}_I/\mathcal{N}(\mathbf{X})$ -bi-measurable.

Proof. We show that ξ is measurable and then that $\xi[C] \in \mathcal{N}(\mathbf{X})$ for any $C \in \mathcal{F}_I$:

- i. A generating family for the σ -algebra $\mathcal{N}(\mathbf{X})$ of subsets of $\mathbf{N}(\mathbf{X})$ is found to be made of subsets of the form

$$C = \{\mu \in \mathbf{N}(\mathbf{X}) \text{ s.t. } \mu(B) = i\},$$

for some $B \in \mathcal{B}(\mathbf{X})$ and some $i \in \mathbb{N}$. The inverse image of C by the mapping ξ is of the form

$$\xi^{-1}[C] = \left\{ f \in \mathbf{F}_I \text{ s.t. } \sum_{x \in \text{dom}(f)} \mathbf{1}_B(f(x)) = i \right\}.$$

To verify that $\xi^{-1}[C] \in \mathcal{F}_I$, we check that

$$(\forall f \in \xi^{-1}[C], \forall f' \in \mathbf{F}_I) \quad f \boldsymbol{\rho}^* f' \Rightarrow f' \in \xi^{-1}[C].$$

By definition we have that

$$f \boldsymbol{\rho}^* f' \Leftrightarrow \exists \nu : \text{dom}(f) \leftrightarrow \text{dom}(f') (f = f' \circ \nu).$$

so that

$$\sum_{x \in \text{dom}(f)} \mathbf{1}_B(f(x)) = \sum_{x \in \text{dom}(f)} \mathbf{1}_B(f'(\nu(x))) = \sum_{x \in \text{dom}(f')} \mathbf{1}_B(f'(x)) = i,$$

and $f' \in \xi^{-1}[C]$ as required.

- ii. To identify a generating family for the σ -algebra \mathcal{F}_I , consider a subset of the form

$$A_{\mathcal{X}} = \{f \in \mathbf{F}_I \text{ s.t. } \text{dom}(f) \supseteq \mathcal{X}, \forall x \in \text{dom}(f) (x \in \mathcal{X} \Leftrightarrow f(x) \in B)\},$$

for some $\mathcal{X} \in \mathcal{X}$ and a Borel subset B of \mathbf{X} , which includes all the functions based on populations having \mathcal{X} as a sub-population that maps the individual in \mathcal{X} into B and all the other individuals outside of B . Then, enlarge the subset $A_{\mathcal{X}}$ by all the functions that are related by $\boldsymbol{\rho}^*$ to any function in it, that is

$$C = \bigcup_{f \in A_{\mathcal{X}}} [f] = \{f \in \mathbf{F}_I \text{ s.t. } \exists X \subseteq \text{dom}(f) (\exists \nu : \mathcal{X} \leftrightarrow X (f \in A_{\nu[\mathcal{X}]}))\}$$

which, denoting $i \doteq |\mathcal{X}|$, can also be expressed as

$$C = \left\{ f \in \mathbf{F}_I \text{ s.t. } \sum_{x \in \text{dom}(f)} \mathbf{1}_B(f(x)) = i \right\}.$$

It follows easily that

$$\xi[C] = \{\mu \in \mathbf{N}(\mathbf{X}) \text{ s.t. } \mu(B) = i\} \in \mathcal{N}(\mathbf{X}).$$

We conclude from [i.](#) and [ii.](#) that ξ is bi-measurable. \square

Theorem [1](#) shows that a stochastic population where all individuals are strongly indistinguishable is essentially equivalent to a point process. To obtain the full equivalence would require to define ξ on $\mathbf{F}_I/\boldsymbol{\rho}^*$, in which case it

would become an isomorphism. This demonstrate that stochastic representation adequately model strongly-indistinguishable populations. Yet, the objective is to be able to represent partially-distinguishable populations and therefore events about specific individuals should also be in the σ -algebra \mathcal{F} . However, considering a random variable on \mathbf{F} , it appears that there is no way of recognising individuals between different realisations, even when these realisations relate to the same population size and structure. In other words, this approach makes all individuals indistinguishable as there would be no way of assessing events based on specific individuals without a means of indexing the distinguished ones.

Example 4. Considering, as in Example 2, a representative set $\mathbf{X} = \{\mathbf{x}, \mathbf{x}'\}$ as a state space, assuming that $\mathcal{X} = \{\{x, x'\} \text{ s.t. } x, x' \in \mathcal{X}_a, x \neq x'\}$, i.e. that population are made of exactly two individuals, and supposing that individuals are always distinguishable, we obtain that

$$\mathbf{F} = \{f : \{x, x'\} \rightarrow \{\mathbf{x}, \mathbf{x}'\} \text{ s.t. } x, x' \in \mathcal{X}_a, f(x) \neq f(x')\}.$$

We can check that $f\rho^*f'$ holds for any $f, f' \in \mathbf{F}$, so that $\rho^* = I$ and \mathbf{F}/ρ^* is a singleton that can be seen as equivalent to the counting measure $\delta_{\mathbf{x}} + \delta_{\mathbf{x}'}$, so that the realisations for the individuals x and x' cannot be distinguished from a random variable on $(\mathbf{F}, \mathcal{F})$.

2.2.2. Second attempt

Since weak indistinguishability is a probabilistic concept, an alternative is to work directly on the set

$$\mathbf{P}_{\mathbf{F}} \doteq \bigcup_{\mathcal{Y} \in \mathcal{Y}} \mathbf{P}(\mathbf{F}_{\mathcal{Y}}),$$

where $\mathbf{P}(\mathbf{E})$ denotes the set of probability measures on a given set \mathbf{E} with its underlying σ -algebra. It is then possible to simplify the set $\mathbf{P}_{\mathbf{F}}$ while preserving the relations of indistinguishability between individuals. For any \mathcal{Y} and \mathcal{Y}' in \mathcal{Y} and any bijection ν between \mathcal{Y} and \mathcal{Y}' , we introduce the mapping $T_{\nu} : \mathbf{F}_{\mathcal{Y}'} \rightarrow \mathbf{F}_{\mathcal{Y}}$ defined by

$$T_{\nu} : f \mapsto f \circ \nu.$$

The mapping defined in (3) can be seen as a special case when $\mathcal{Y} = \mathcal{Y}'$.

Definition 5. For any populations $\mathcal{Y}, \mathcal{Y}' \in \mathcal{Y}$, any $P \in \mathbf{P}(\mathbf{F}_{\mathcal{Y}})$ and any $P' \in \mathbf{P}(\mathbf{F}_{\mathcal{Y}'})$, let the binary relation ρ on $\mathbf{P}_{\mathbf{F}}$ be defined as

$$P\rho P' \Leftrightarrow \exists \nu : \mathcal{Y} \xrightarrow{\sim} \mathcal{Y}' (P = (T_{\nu})_* P'). \quad (6)$$

Since each probability measure P in $\mathbf{P}_{\mathbf{F}}$ is defined on a single population in \mathcal{Y} , the latter can be recovered and will be denoted \mathcal{Y}_P or (\mathcal{X}_P, τ_P) . If individuals are independent under a given probability measure $P \in \mathbf{P}_{\mathbf{F}}$ then the equivalence class $[P]$ of probability measures related to P via ρ is found to be

$$[P] = \{P' \text{ s.t. } \exists \nu : \mathcal{Y}_P \xrightarrow{\sim} \mathcal{Y}_{P'} (\forall x \in \mathcal{X}_P (p_x = p'_{\nu(x)}))\}.$$

This result highlights the structure of the equivalence relation ρ and of the mapping T_ν in (6).

Note that the definition of ρ does not depend on the relation η of weak indistinguishability. Indeed, weak indistinguishability is more an observed property of a representation rather than a building block that would impose some sort of structure on the mathematical construction of it.

A given equivalence class in $\mathbf{P}_\mathbf{F}/\rho$ allows for describing the randomness of a population of a given size and structure without knowing the actual population state in \mathcal{X} as required. Such an equivalence class is referred to as a *population representation* or simply as a *representation*, and when individuals are independent, the induced probability measures on \mathbf{X} are called *individual representations*.

Example 5. Considering again the case of Example 4 it follows that

$$\mathbf{P}_\mathbf{F} = \{P \in \mathbf{P}(\mathbf{F}_{(\{x,y\}, O)}) \text{ s.t. } x, y \in \mathcal{X}_a\}.$$

Focusing on the subset $\mathbf{P}_\mathbf{F}^*$ of $\mathbf{P}_\mathbf{F}$ for which individuals are weakly distinguishable and independent, for the sake of simplicity, we find that

$$P\rho P' \Leftrightarrow \exists \nu : \{x, y\} \leftrightarrow \{x', y'\} ((p_x = p'_{\nu(x)}) \wedge (p_y = p'_{\nu(y)})),$$

for any $P, P' \in \mathbf{P}_\mathbf{F}^*$, where $\mathcal{X}_P = \{x, y\}$ and $\mathcal{X}_{P'} = \{x', y'\}$. In this setup, a point in $\mathbf{P}_\mathbf{F}^*/\rho$, which is in fact an equivalence class, correspond to the configuration where the uncertainty about one individual is described by a given law p and the uncertainty about the other individual is described by a given law p' , the individuals being weakly indistinguishable if $p = p'$. In other words, individuals are labeled by the probability measures describing the uncertainty about them, these labels being shared by indistinguishable individuals by definition.

The set $\mathbf{P}_\mathbf{F}/\rho$ is not however a full answer to the question of the representation of populations since elements of it correspond to a given size and a given structure, i.e. a given type of strong indistinguishability. Yet, the size and structure of a population are generally unknown and possibly random, and there might be second-order uncertainties on the probability measures in $\mathbf{P}_\mathbf{F}$ themselves. Indeed, in general, there are many possible distinct configurations for each given cardinality and structure, which can only be represented by random probability measures for each choice of $[\mathcal{Y}] \in \mathcal{Y}/\approx$. The set $\mathbf{P}_\mathbf{F}$ also has to be embedded with a suitable σ -algebra: even when a topology on \mathcal{Y} is available, the corresponding topology on $\mathbf{P}_\mathbf{F}$ would not be suitable for our purpose since it would allow for distinguishing representations based on a given population $\mathcal{X} \in \mathcal{X}$. Instead, we consider the initial topology induced by the quotient map of ρ and we denote $\mathcal{P}_\mathbf{F}$ the corresponding Borel σ -algebra. There is no natural reference measure on $\mathcal{P}_\mathbf{F}$, but we assume that such a measure is given case by case via a countable subset or a parametric family of probability measures. Similarly, the σ -algebra on \mathcal{Y} is assumed to be induced by the discrete topology on \mathcal{Y}/\approx .

A random variable \mathfrak{M} on $(\mathbf{P}_\mathbf{F}, \mathcal{P}_\mathbf{F})$ describes all the uncertainties about the system of interest and is referred to as a *stochastic representation*. The

interpretation of \mathfrak{M} can be made easier by separating its law $P \doteq \mathfrak{M}_* \mathbb{P}$ into a marginal and a conditional as

$$(\forall B \in \mathcal{P}_{\mathbf{F}}) \quad P(B) = \mathbb{E}[P(B | \mathfrak{Y})],$$

where \mathfrak{Y} is the random population induced by \mathfrak{M} on \mathcal{Y} , and $P(\cdot | \mathfrak{Y})$ is a version of the conditional law of \mathfrak{M} given \mathfrak{Y} , i.e. the law representing the second-order uncertainties given the size and structure of the population. This separation of the randomness is straightforward but helps to interpret the behaviour of \mathfrak{M} : first a size and a structure $[\mathcal{Y}] \in \mathcal{Y}/\approx$ is randomly selected for the population, then a probability measure on $\mathbf{F}_{\mathcal{Y}}$ is drawn, where \mathcal{Y} is any element of $[\mathcal{Y}]$, describing the uncertainty about the considered type of population and ensuring that there is no specific knowledge about strongly indistinguishable individuals. Which population has been chosen from $[\mathcal{Y}]$ is irrelevant since the mapping $P(B | \mathfrak{Y})$ has to be measurable for any $B \in \mathcal{P}_{\mathbf{F}}$.

Remark 1. The probability $\mathbb{P}(\mathfrak{Y} \in [(\mathcal{X}, \tau)])$ only depends on the size of \mathcal{X} and on the size of the subsets in \mathcal{X}/τ . For instance, we can evaluate the probability for a realisation \mathcal{Y} of \mathfrak{Y} to contain exactly 3 strongly indistinguishable individuals and 2 weakly distinguishable ones, however, we cannot assess the probability of any event regarding the states of these individuals in \mathcal{X}_a . Likewise, $P_{\mathfrak{M}|\mathfrak{Y}}$ only allows for evaluating the probability of events about some individuals being represented by some probability measures, for instance, for the 3 strongly indistinguishable individuals to be independent and associated with the individual law $p \in \mathbf{P}(\mathbf{X})$ and for the 2 weakly distinguishable individuals to be dependent and associated with a joint law (that will be non symmetrical if they have been distinguished).

2.3. Statistics

The stochastic representation \mathfrak{M} is a random element of $\mathbf{P}_{\mathbf{F}}$ inducing a random size and a random structure via the randomly selected law. Because the realisations of \mathfrak{M} are probability measures on different spaces, these realisations are not directly summable, yet statistics for some aspects of \mathfrak{M} can be defined. In particular, the n^{th} -order moment for the number of individuals, if it exists, is equal to $\mathbb{E}[|\mathcal{X}_{\mathfrak{M}}|^n]$. Similarly, equivalence relations are not summable, yet the n^{th} -order moments for the number of strongly indistinguishable sub-populations is found to be equal to $\mathbb{E}[|\mathcal{X}_{\mathfrak{M}}/\tau_{\mathfrak{M}}|^n]$ and is finite by construction.

By definition, for any $\mathcal{Y} = (\mathcal{X}, \tau) \in \mathcal{Y}$, the measurable space $(\mathbf{F}_{\mathcal{Y}}, \mathcal{F}_{\mathcal{Y}})$ excludes events regarding a subset of individuals specifically when this subset does not form a union of elements of \mathcal{X}/τ . Henceforth, the set $\mathbf{N}(\mathbf{X})$ of integer-valued measures is equipped with its Borel σ -algebra $\mathcal{N}(\mathbf{X})$.

Proposition 3. *Let $\mathcal{Y} = (\mathcal{X}, \tau) \in \mathcal{Y}$ and let X be a subset of \mathcal{X} , then the mapping T_X defined as*

$$T_X : \mathbf{F}_{\mathcal{Y}} \rightarrow \mathbf{N}(\mathbf{X}) \tag{7a}$$

$$f \mapsto \sum_{x \in X} \delta_{f(x)}, \tag{7b}$$

is $\mathcal{F}_{\mathcal{Y}}/\mathcal{N}(\mathbf{X})$ -measurable if and only if X is an element of a partition that is coarser than \mathcal{X}/τ .

Proof. As mentioned before, the σ -algebra $\mathcal{N}(\mathbf{X})$ is generated by subsets of the form $C = \{\mu \in \mathbf{N}(\mathbf{X}) \text{ s.t. } \mu(B) = i\}$, for some $B \in \mathcal{B}(\mathbf{X})$ and some $i \in \mathbb{N}$, and it holds that

$$T_X^{-1}(C) = \left\{ f \in \mathbf{F}_{\mathcal{Y}} \text{ s.t. } \sum_{x \in X} \mathbf{1}_B(f(x)) = i \right\}.$$

The mapping T_X is measurable if and only if

$$(\forall f \in T_X^{-1}[C], \forall f' \in \mathbf{F}_{\mathcal{Y}}) \quad f \rho f' \Rightarrow f' \in T_X^{-1}[C],$$

which is equivalent to

$$(\forall f \in T_X^{-1}[C], \forall \sigma \in \text{Sym}(\mathcal{X}, \tau)) \quad \sum_{x \in X} \mathbf{1}_B(f(\sigma(x))) = i.$$

This last statement holds if and only if $\sigma[X] = X$ for all $\sigma \in \text{Sym}(\mathcal{X}, \tau)$, i.e. if and only if there exists partition of \mathcal{X} containing X and being coarser than \mathcal{X}/τ . \square

One can then study the law $(T_X)_*P$ for any $P \in \mathbf{P}(\mathbf{F}_{\mathcal{Y}})$, which is a point-process distribution. The mapping T_X does not retain information about the structure $\tau|_X$ of the subset X but distinguishability will be lost anyway when considering expectation over \mathfrak{M} . For instance, the mean number of induced point-process laws within a measurable subset B of $\mathbf{P}(\mathbf{N}(\mathbf{X}))$ given by indistinguishable sub-populations of fixed size $m \in \mathbb{N}$ can be expressed as $\mathbb{E}[\chi_B^m(P)]$ with

$$\chi_B^m(P) \doteq \sum_{\substack{X \in \mathcal{X}_P/\tau_P \\ |X|=m}} \mathbf{1}_B((T_X)_*P).$$

This is only an example, statistics about sub-populations can be defined based on the characteristics of the induced point-process laws or on the characteristics of the induced point processes themselves, by studying moments corresponding to the probability mass induced by \mathfrak{M} on given subsets of $\mathbf{N}(\mathbf{X})$.

Example 6. In the case where all individuals are weakly distinguishable via \mathfrak{M} , that is when $\tau_{\mathfrak{M}} = O$ holds almost surely, other summable quantities that are induced by the stochastic representation \mathfrak{M} are the number of marginal individual laws in a given measurable subset B of $\mathbf{P}(\mathbf{X})$ as well as the marginal individual laws themselves, since they are defined on the same space. In this specific case, the mapping defined in (7) can be simplified as follows: for any $\mathcal{Y} = (\mathcal{X}, \tau) \in \mathcal{Y}$ such that $\tau = O$, a measurable mapping T_x on $\mathbf{F}_{\mathcal{Y}}$ is introduced as

$$\begin{aligned} T_x : \mathbf{F}_{\mathcal{Y}} &\rightarrow \mathbf{X} \\ f &\mapsto f(x) \end{aligned}$$

defined for any $x \in \mathcal{X}$. This mapping can be seen as a special case of (7) with $X = \{x\}$ and $f(x)$ being used directly instead of $\delta_{f(x)}$. Then, for a given $P \in \mathbf{P}_{\mathbf{F}}$ such that $\tau_P = O$ and a given $x \in \mathcal{X}_P$, the marginal individual law of x by P is the pushforward $(T_x)_*P$. The quantities of interest are then defined for any measurable subset B of $\mathbf{P}(\mathbf{X})$ and any measurable subset B' of \mathbf{X} as

$$\begin{aligned}\chi_B(P) &\doteq \sum_{x \in \mathcal{X}_P} \mathbf{1}_B((T_x)_*P), \\ \bar{\chi}_{B'}(P) &\doteq \sum_{x \in \mathcal{X}_P} (T_x)_*P(B').\end{aligned}$$

The corresponding n^{th} -order moments, if they exist, are thus $\mathbb{E}[\chi_B(\mathfrak{M})^n]$ and $\mathbb{E}[\bar{\chi}_{B'}(\mathfrak{M})^n]$. For instance, $\mathbb{E}[\chi_B(\mathfrak{M})]$ is the expected number of individuals with marginal law within B while $\mathbb{E}[\bar{\chi}_{B'}(\mathfrak{M})]$ is the probability mass given by the average marginal individual law to the subset B' .

Many applications are concerned with the study of populations where the individuals are independent. The simplifications induced by such an assumption are important enough to justify studying this case specifically, and so is done in the next section.

3. Alternative formulation

The objective is now to show that the problem can be formulated on more standard sets than $\mathbf{P}_{\mathbf{F}}$. We focus on one alternative formulation which relies on integer-valued measures, however, other formulations are possible, e.g. with product measures on suitably defined spaces. These types of formulation already exist for point processes as described in [20] and [22]. The following assumption will henceforth be considered:

A.1 Individuals are independent.

The subset of $\mathbf{P}_{\mathbf{F}}$ composed of probability measures for which all individuals are independent is denoted $\mathbf{P}_{\mathbf{F}}^*$ and is equipped with the σ -algebra $\mathcal{P}_{\mathbf{F}}^*$ induced by $\mathcal{P}_{\mathbf{F}}$. For a given $P \in \mathbf{P}_{\mathbf{F}}^*$, we denote $\mathcal{X}_P \in \mathcal{X}$ the population on which P is based and $\{p_x\}_{x \in \mathcal{X}_P}$ the corresponding family of individual laws on \mathbf{X} .

One of the most direct alternative formulations uses the concept of integer-valued measures or counting measures. A connection between the specific notion of population representation and the more common concept of counting measure is established in the following proposition. Since $\mathbf{P}(\mathbf{X})$ is a Polish space when equipped with the topology induced by the Prokhorov metric [23], the set $\mathbf{N}(\mathbf{P}(\mathbf{X}))$ can also be made Polish [19] and is therefore equipped with its Borel σ -algebra denoted $\mathcal{N}(\mathbf{P}(\mathbf{X}))$. Also, the Borel σ -algebra of $\mathbf{P}(\mathbf{X})$ is denoted by $\mathcal{P}(\mathbf{X})$.

Theorem 2. *The mapping $\zeta : \mathbf{P}_{\mathbf{F}}^* \rightarrow \mathbf{N}(\mathbf{P}(\mathbf{X}))$, defined as*

$$\zeta : P \mapsto \sum_{x \in \mathcal{X}_P} \delta_{p_x}, \quad (10)$$

is $\mathcal{P}_{\mathbf{F}}^*/\mathcal{N}(\mathbf{P}(\mathbf{X}))$ -measurable.

Proof. The Borel σ -algebra on $\mathbf{N}(\mathbf{P}(\mathbf{X}))$ is the one generated by subsets of the form

$$C = \{\mu \in \mathbf{N}(\mathbf{P}(\mathbf{X})) \text{ s.t. } \mu(B) = i\},$$

for some $B \in \mathcal{B}(\mathbf{P}(\mathbf{X}))$ and $i \in \mathbb{N}$. The inverse image of C by ζ is found to be

$$\zeta^{-1}[C] = \left\{ P \in \mathbf{P}_{\mathbf{F}} \text{ s.t. } \sum_{x \in \mathcal{X}_P} \mathbf{1}_B(p_x) = i \right\},$$

where \mathcal{X}_P is the population on which P is defined and $\{p_x\}_{x \in \mathcal{X}_P}$ is the indexed family of probability measures on \mathbf{X} induced by P . Following the same route as in the proof of Theorem 1, we can verify that $\zeta^{-1}[C] \in \mathcal{P}_{\mathbf{F}}^*$. \square

Theorem 2 shows stochastic representations can be expressed as a random counting measures, or point process, on the set of probability measures on \mathbf{X} , and such will be the understanding in this section. The transformation ζ introduced in this proposition does not preserve the representation of strong indistinguishability and is not bi-measurable as a consequence. This can be seen as beneficial in practice since the observability of strong indistinguishability is often out of reach. The only individuals that are known to be strongly indistinguishable in this case are the ones that are almost surely at the same point of the state space, i.e., the ones which law is known to be of the form $\delta_{\mathbf{x}}$ for some $\mathbf{x} \in \mathbf{X}$.

Remark 2. It is possible to relax Assumption A.1 to: individuals that are not strongly indistinguishable are independent. In this case, the corresponding subset of stochastic representations could be mapped to $\mathbf{N}(\mathbf{P}(\mathbf{X}^\times))$, with the set \mathbf{X}^\times defined as

$$\mathbf{X}^\times \doteq \{\psi_\infty\} \cup \bigcup_{k \geq 1} \mathbf{X}^k, \quad (11)$$

where the point state denoted ψ_∞ represents the case where infinitely many individuals are at point ψ . In this configuration, the relation of strong indistinguishability can be preserved, but at the expense of a more complex set of counting measures.

As a point process, \mathfrak{M} can be characterised by its probability-generating functional (p.g.fl.) G , defined for any non-negative bounded measurable function h on $\mathbf{P}(\mathbf{X})$ as [24]

$$\begin{aligned} G(h) &\doteq \mathbb{E} \left[\exp \left(\int \log h(p) \mathfrak{M}(dp) \right) \right] \\ &= p(0) + \sum_{n \geq 1} p(n) \int \prod_{i=1}^n h(p_i) P(d(p_1, \dots, p_n) | n), \end{aligned}$$

where $p \in \mathbf{P}(\mathbb{N})$ is defined as $p(n) = \mathbb{P}(\mathfrak{M}(\mathbf{P}(\mathbf{X})) = n)$ for any $n \geq 0$ and $P(\cdot | n)$ is the distribution of \mathfrak{M} on $\mathbf{P}(\mathbf{X})^n$ conditioned on $\mathfrak{M}(\mathbf{P}(\mathbf{X})) = n$, for any $n > 0$.

The simplicity of this integer-valued measure formulation comes from the fact that the state space does not actually appear in the equations, allowing for more flexibility in the expressed quantity. This formulation has been used in [25] in the context of Bayesian data assimilation for multi-object systems. Yet, it is sometimes necessary to assess events for the stochastic population at the level of the individual state space \mathbf{X} . A second alternative would be to express the stochastic representation on a product space based on \mathbf{X} , by collapsing the two levels of probabilistic structures considered so far. This formulation has been used for expressing Bayesian data-assimilation algorithms in [16, 17] and for deriving approximate solutions [26] used in [15, 18].

3.1. Parametrised family of probability measures

A special case of interest is found when the support of the considered stochastic representations is within a family of probability measures parametrised by a set $\Theta \subseteq \mathbb{R}^{d_\Theta}$ for some $d_\Theta > 0$. This enables some of the properties of stochastic representations to be studied on the simpler set Θ . The following additional assumption is henceforth considered:

A.2 Stochastic representations take values in a parametrised family of probability measures.

Under Assumption A.2, let $\mathcal{S}_\Theta = \{p_\theta\}_{\theta \in \Theta}$ be an identifiable family of probability measures on \mathbf{X} encompassing the support of \mathfrak{M} . In this context, identifiability means that $p_\theta \neq p_{\theta'}$ whenever the parameters $\theta, \theta' \in \Theta$ are different. The point process \mathfrak{M} induces a point process \mathfrak{N} on Θ in the following way:

$$\mathfrak{N} = (F^{-1})_* \mathfrak{M}$$

where $F : \Theta \ni \theta \mapsto p_\theta \in \mathbf{P}(\mathbf{X})$ is assumed to be bi-measurable. Straightforwardly, any point process \mathfrak{N}' on Θ induces a point process on $\mathbf{P}(\mathbf{X})$ defined as $F_* \mathfrak{N}'$. One of the consequences on this relation is the ability to recover statistics for \mathfrak{M} from the ones for \mathfrak{N} , for instance the n^{th} -order moment evaluated at the measurable subset B of $\mathbf{N}(\mathbf{P}(\mathbf{X}))$ can be recovered via

$$\mathbb{E}[\mathfrak{M}(B)^n] = \mathbb{E}[\mathfrak{N}(F^{-1}(B))^n].$$

The p.g.fl. of \mathfrak{M} can now be equivalently expressed as

$$\begin{aligned} G(h) &= \mathbb{E} \left[\exp \left(\int \log h(p_\theta) \mathfrak{N}(d\theta) \right) \right] \\ &= p(0) + \sum_{n \geq 1} p(n) \int \prod_{i=1}^n h(p_{\theta_i}) Q(d(\theta_1, \dots, \theta_n) | n), \end{aligned}$$

where $Q(\cdot | n)$ is the distribution of \mathfrak{N} on Θ^n conditioned on $\mathfrak{N}(\Theta) = n$, for any $n > 0$. If the population under consideration is fully distinguishable almost surely then the point process \mathfrak{M} is simple and Q admits a density w.r.t. the Lebesgue measure on Θ .

3.2. Discrete set of probability measures

We also formulate an assumption that is of interest when devising practical estimation algorithms:

A.3 The set Θ is countable.

As a consequence of Assumption **A.3**, the point process \mathfrak{N} induced by \mathfrak{M} is equivalent to a random variable \mathbf{N} on the set $\bar{\mathbb{N}}^\Theta$, with $\bar{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$. Then \mathfrak{M} can be expressed as

$$\mathfrak{M} = \sum_{\theta \in \Theta} N_\theta \delta_{p_\theta}.$$

Note that \mathbf{N} verifies $N_\theta < \infty$ for any $\theta \in \Theta$ such that $p_\theta \neq \delta_\psi$. A realisation μ of \mathfrak{M} can be denoted $\mu_{\mathbf{n}}$ with \mathbf{n} the corresponding realisation of \mathbf{N} in order to underline the multiplicity of each atom in \mathcal{S}_Θ . The law P of \mathfrak{M} on $\mathbf{N}(\mathbf{P}(\mathbf{X}))$ can then be expressed as

$$P(B) = \int \mathbf{c}(\mathrm{d}\mathbf{n}) \mathbf{1}_B(\mu_{\mathbf{n}})$$

for any Borel subset B of $\mathbf{N}(\mathbf{P}(\mathbf{X}))$, where \mathbf{c} is the induced probability measure on $\bar{\mathbb{N}}^\Theta$. This way of representing stochastic populations is useful when performing filtering [27, Chapt. 3] since finite collections of individual representations are often available in practice, so that Assumption **A.3** is verified.

Example 7. If a population is known to contain exactly 3 individuals and if the only available individual representations for these individuals are the ones in the set $\mathcal{S}_\Theta = \{p_1, p_2\}$, in which case $\Theta = \{1, 2\}$ and elements of $\bar{\mathbb{N}}^\Theta$ can be seen as pairs of integers, then the population representation can be any of the following:

$$\mu_{3,0} = 3\delta_{p_1}, \quad \mu_{2,1} = 2\delta_{p_1} + \delta_{p_2}, \quad \mu_{1,2} = \delta_{p_1} + 2\delta_{p_2}, \quad \mu_{0,3} = 3\delta_{p_2}.$$

For instance, $\mu_{2,1}$ describes the case where the uncertainty about two of the individuals is described by p_1 , so that these two individuals are indistinguishable, and the uncertainty about the other individual is described by p_2 . In this form it is not known whether the two weakly indistinguishable individuals are also strongly indistinguishable or not. Any corresponding stochastic representation \mathfrak{M} is a point process on $\mathbf{P}(\mathbf{X})$ verifying

$$\mathfrak{M}(\mathbf{P}(\mathbf{X}) - \{p_1, p_2\}) = 0 \quad a.s.,$$

so that \mathfrak{M} can be simply described by the multiplicities it assigns to probability measures in \mathcal{S}_Θ .

Identifying a countable family \mathcal{S}_Θ of probability measure and additionally assuming that $N_\theta < \infty$ even if $p_\theta = \delta_\psi$ enables a simplification of the expression of the p.g.fl. of \mathfrak{M} to

$$\begin{aligned} G(h) &= \mathbb{E} \left[\prod_{\theta \in \Theta} h(p_\theta)^{N_\theta} \right] \\ &= \sum_{\mathbf{n} \in \bar{\mathbb{N}}^\Theta} \mathbf{c}(\mathbf{n}) \prod_{\theta \in \Theta} h(p_\theta)^{n_\theta}, \end{aligned}$$

which is related to the probability-generating functional of \mathfrak{N} as expected. For instance, if $\Theta = \{1, \dots, k\}$, then $G(h) = G'(h(p_1), \dots, h(p_k))$, with G' the probability-generating function of \mathbf{N} defined as

$$G'(z_1, \dots, z_k) \doteq \sum_{\mathbf{n} \in \mathbb{N}^\Theta} \mathbf{c}(\mathbf{n}) z_1^{\mathbf{n}_1} \dots z_k^{\mathbf{n}_k}.$$

Assumption [A.3](#) also yields a simpler expression of the statistics induced by a stochastic representation \mathfrak{M} on $\mathbf{N}(\mathbf{P}(\mathbf{X}))$. Of particular interest are the mean $M(B)$ and variance $V(B)$ for the number of individual laws within a measurable subset B of $\mathbf{P}(\mathbf{X})$, characterised by

$$\begin{aligned} M(B) &\doteq \mathbb{E}[\mathfrak{M}(B)], \\ V(B) &\doteq \mathbb{E}[\mathfrak{M}(B)^2] - M(B)^2, \end{aligned}$$

whenever they exist. These quantities are well defined since \mathfrak{M} is a random measure. If the quantities of interest are the mean and variance on the state space \mathbf{X} , then the mapping

$$\begin{aligned} \Phi_B : \mathbf{P}(\mathbf{X}) &\rightarrow \mathbb{R} \\ p &\mapsto p(B), \end{aligned}$$

can be introduced for any $B \in \mathcal{B}(\mathbf{X})$ and is $\mathcal{P}(\mathbf{X})/\mathcal{B}(\mathbb{R})$ -measurable by [\[19, Proposition A2.5.IV\]](#). The *collapsed* first moment $M'(B')$ and variance $V'(B')$, describing the number of individuals within $B' \in \mathcal{B}(\mathbf{X})$ can then be defined as [\[25\]](#)

$$\begin{aligned} M'(B) &\doteq \mathbb{E}[\mathfrak{M}(\Phi_B)] = \sum_{\theta \in \Theta} p_\theta(B), \\ V'(B) &\doteq \mathbb{E}[\mathfrak{M}(\Phi_B)^2] - M'(B)^2 = \sum_{\theta, \theta' \in \Theta} \text{cov}_{\theta, \theta'} p_\theta(B) p_{\theta'}(B). \end{aligned}$$

where $m_\theta \doteq \mathbb{E}[\mathbf{N}_\theta]$ and $\text{cov}_{\theta, \theta'} \doteq \mathbb{E}[\mathbf{N}_\theta \mathbf{N}_{\theta'}] - \mathbb{E}[\mathbf{N}_\theta] \mathbb{E}[\mathbf{N}_{\theta'}]$. These relations between \mathbf{N} and \mathfrak{M} are connected to the relation between the p.g.fl. G and the probability-generating function G' . Even in the simple configuration induced by Assumption [A.3](#), the structure of the proposed representation of stochastic populations enables more diverse types of statistics to be computed when compared to point processes on the state space, which is practically relevant for describing filtering algorithms for multi-object dynamical systems [\[28\]](#).

Conclusion

Starting from general considerations about the concepts of individual and population and about the partially-indistinguishable knowledge that may be available about them, we went across increasingly general notions in an attempt to faithfully describe the multi-faceted nature of the corresponding uncertainties. After a suitable level of generality was reached, an alternative way of

expressing the uncertainty about these complex systems has been introduced. This alternative expression highlights the nature of the proposed representation by identifying it with a point process on the set of probability measures on the individual state space, under the assumption of independence between individuals.

References

- [1] N. Chenouard, et al., Objective comparison of particle tracking methods, *Nature methods* 11 (3) (2014) 281.
- [2] J. Mullane, B.-N. Vo, M. D. Adams, B.-T. Vo, A random-finite-set approach to Bayesian SLAM, *IEEE Transactions on Robotics* 27 (2) (2011) 268–282.
- [3] K. Okuma, A. Taleghani, N. De Freitas, J. J. Little, D. G. Lowe, A boosted particle filter: Multitarget detection and tracking, in: *Computer Vision-ECCV 2004*, Springer, 2004, pp. 28–39.
- [4] J. Hofbauer, K. Sigmund, *Evolutionary games and population dynamics*, Cambridge university press, 1998.
- [5] P. Turchin, *Complex population dynamics: a theoretical/empirical synthesis*, Vol. 35, Princeton University Press, 2003.
- [6] C. J. Geyer, J. Møller, Simulation procedures and likelihood inference for spatial point processes, *Scandinavian journal of statistics* (1994) 359–373.
- [7] P. J. Green, Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika* 82 (4) (1995) 711–732.
- [8] S. S. Blackman, *Multiple-target tracking with radar applications*, Vol. 1, Artech House, Norwood, 1986.
- [9] R. P. S. Mahler, *Statistical multisource-multitarget information fusion*, Artech House, Boston, 2007.
- [10] F. Caron, P. Del Moral, A. Doucet, M. Pace, On the conditional distributions of spatial point processes, *Advances in Applied Probability* 43 (2) (2011) 301–307.
- [11] M. Pace, P. Del Moral, Mean-field PHD filters based on generalized Feynman-Kac flow, *IEEE Journal of Selected Topics in Signal Processing* 7 (3) (2013) 484–495.
- [12] P. Del Moral, J. Houssineau, Particle association measures and multiple target tracking, in: *Theoretical Aspects of Spatial-Temporal Modeling*, Springer, 2015, pp. 1–30.
- [13] P. Del Moral, *Mean field simulation for Monte Carlo integration*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, 2013.

- [14] J. Houssineau, D. E. Clark, Bayesian data assimilation based on a family of outer measures, arXiv preprint arXiv:1611.02989.
- [15] Y. Pailhas, J. Houssineau, Y. Petillot, D. E. Clark, Tracking with MIMO sonar systems: applications to harbour surveillance, Accepted to: IET Radar, Sonar & Navigation.
- [16] E. Delande, J. Houssineau, D. E. Clark, Performance metric in closed-loop sensor management for stochastic populations, in: IEEE Sensor Signal Processing for Defence (SSPD), 2014.
- [17] E. Delande, C. Frueh, J. Franco, J. Houssineau, D. E. Clark, Multi-object filtering for space situational awareness, Submitted to: Journal of Guidance, Control, and Dynamics.
- [18] J. Houssineau, D. E. Clark, P. Del Moral, A sequential Monte Carlo approximation of the HISP filter, in: European Signal Processing Conference (EUSIPCO), 2015.
- [19] D. Daley, D. Vere-Jones, An Introduction to The Theory of Point Processes, 2nd Edition, Springer, 2003.
- [20] J. Moyal, The general theory of stochastic population processes, Acta Mathematica (1962) 1–31.
- [21] O. Macchi, The coincidence approach to stochastic point processes, Advances in Applied Probability (1975) 83–122.
- [22] K. Itô, Stochastic processes: lectures given at Aarhus University, Springer Science & Business Media, 2013.
- [23] Y. V. Prokhorov, Convergence of random processes and limit theorems in probability theory, Theory of Probability & Its Applications 1 (2) (1956) 157–214.
- [24] D. Daley, D. Vere-Jones, An Introduction to The Theory of Point Processes, Vol. II, 2nd Edition, Springer, 2008.
- [25] E. Delande, J. Houssineau, D. E. Clark, Multi-object filtering with stochastic populations, ArXiv preprint arXiv:1501.04671.
- [26] J. Houssineau, D. E. Clark, Multi-target filtering with linearised complexity, ArXiv preprint arXiv:1404.7408v2.
- [27] J. Houssineau, Representation and estimation of stochastic populations, Ph.D. thesis, Heriot-Watt University (2015).
- [28] E. Delande, M. Uney, J. Houssineau, D. E. Clark, Regional variance for multi-object filtering, IEEE Transactions on Signal Processing 62 (13) (2014) 3415–3428.